UILU-ENG 78 1712

USE OF THE SINGULAR VALUE DECOMPOSITION TO INCREASE EXECUTION AND STORAGE EFFICIENCY OF THE MANTEUFFEL ALGORITHM FOR THE SOLUTION OF NONSYMMETRIC LINEAR· SYSTEMS

by

Paul E. Saylor

April 1978

DEPARTMENT OF COMPUTER SCIENCE
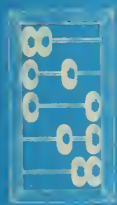**UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN · URBANA, ILLINOIS**

USE OF THE SINGULAR VALUE DECOMPOSITION TO INCREASE
EXECUTION AND STORAGE EFFICIENCY OF THE MANTEUFFEL
ALGORITHM FOR THE SOLUTION OF NONSYMMETRIC LINEAR SYSTEMS

by

Paul E. Saylor

April 1978

Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

Introduction. The main concern here is the efficient use of the Manteuffel algorithm [9] for computing Chebyshev parameters to accelerate the iterative solution of real nonsymmetric linear systems. There are two objectives. One is to determine when to call the Manteuffel algorithm; the other is to reduce the costs of execution and storage. A consequence of knowing when to call the algorithm is that the iteration takes fewer steps. Apart from these, it is also shown that the Manteuffel algorithm may be used with the Jacobi semi-iterative (JSI) method as well as Richardson's method for which it was originally developed.

Chebyshev acceleration parameters are computed from the convex hull of eigenvalues of the matrix associated with the linear system. Eigenvalues are obtained from the power method, which, for an arbitrary square matrix B with a single dominant eigenvalue, is based on the observation that the Krylov sequence $y, \ldots, B^k y$ tends to that eigenvector corresponding to the dominant eigenvalue. Precisely one nonreal eigenvalue of a real matrix cannot be dominant because eigenvalues and the corresponding eigenvectors appear in conjugate pairs. A dominant nonreal eigenvalue would cause $B^k y$ to be a linear combination of two eigenvectors, one the conjugate of the other. If so, $B^k y$, $B^{k+1} y$ and $B^{k+2} y$ are linearly dependent, for large enough k. The coefficients expressing linear dependence,

(1)
$$c_0 B^k y + c_1 B^{k+1} y = -c_2 B^{k+2} y,$$

are seen to be the coefficients of a polynominal whose roots are the dominant eigenvalue and its conjugate. A least squares solution of the overdetermined system (1) yields c_0, c_1, and c_2 [13, p. 580]. The details are in Section (5).

The computation is less expensive and takes less storage if fewer, randomly selected components of $B^k y$, $B^{k+1} y$, $B^{k+2} y$ are used, which is equivalent to replacing $B^k y$, $B^{k+1} y$, $B^{k+2} y$ in the overdetermined system by their projections, $P(B^k y)$, $P(B^{k+1} y)$, $P(B^{k+2} y)$ onto a subspace. The projections must also be linearly dependent.

The linear dependence of a set of vectors, $\{v_1, \ldots, v_k\}$, is equivalent to the existence of a zero singular value of the matrix $(v_1 \ldots v_k)$. The question of whether $P(B^k y)$, $P(B^{k+1} y)$ and $P(B^{k+2} y)$ yield satisfactory eigenvalues reduces to a test of the smallest (in magnitude) singular value of the matrix, D, with these column vectors. Other tests are possible. As an alternative one could determine the rank of $D^T D$ (suggested by Manteuffel [10, p. 19] for a slightly different application) but an efficient, stable algorithm due to Businger-Golub [1] is available to compute singular values and is generally superior [10, p. 382]. If the smallest singular value is too large, that is if the column vectors of

$$(2) \qquad\qquad D = (P(B^k y), P(B^{k+1} y), P(B^{k+2} y))$$

are independent, there is no information in the singular values to suggest whether it is because k should have been larger or because the projection was onto a badly chosen subspace. The smallest singular value is like clinical temperature, which may show a patient is sick, but gives no diagnosis.

In practice, rather than test the smallest (in magnitude) singular value to determine linear dependence, the ratio of the smallest (in magnitude) to the largest (in magnitude) should be tested. If there is linear dependence, the next step is to solve the least squares problem, which may be done at no additional cost by using some of the results from the computation of the singular values (details in section 8).

Current implementations of the Manteuffel algorithm compute new eigenvalues after a fixed number of steps of the power method, a costly strategy if the eigenvector estimates are mature after fewer steps. Linear dependence should be monitored in order to compute new eigenvalues precisely when they would be accurate. If the number of rows of D is small, it would be economical to compute singular values to do this after each iterative step so new eigenvalues could be computed precisely when they are available, resulting in an efficient iteration. Computing singular values to monitor linear dependence thus yields two benefits: lower overhead in executing the Manteuffel algorithm; and fewer iterative steps.

The basic facts about Richardson's method are presented in section 2, with the techniques to estimate iteration parameters dynamically (the Manteuffel algorithm) explained briefly in section 3. In section 4, it is shown how to apply the Manteuffel algorithm to the Jacobi semi-iterative method and to a generalization of Richardson's method which includes the JSI method. Section 5 is a discussion of the power method for nonreal eigenvalues. A simple way to reduce the labor by working with fewer components, recommended by Wrigley [14], is explained in the next section. Whether fewer components will be effective reduces to a study of linear dependence by using the singular values of matrix D, above in (2) (section 7). The computational work of the singular value decomposition and how it applies to computing eigenvalues are given in section 8.

2. <u>Richardson's Method</u>. The basic iteration is a gradient procedure

$$(3) \qquad x^{(n+1)} = x^{(n)} - t_n (Ax^{(n)} - b).$$

The true error, defined to be $e^{(n)} = x - x^{(n)}$, satisfies

(4)
$$e^{(n)} = (I - t_{n-1}A)e^{(n-1)} = R_n(A)e^{(0)}$$

where

$$R_n(\lambda) = (1 - t_{n-1}\lambda)\ldots(1 - t_0\lambda),$$

called the error propagator polynomial, is such that

(5)
$$R_n(0) = 1.$$

For A nonsymmetric with eigenvalues in, say, the right half plane, the gradient procedure converges if the error propagator is

$$P_n(\lambda) = T_n[(d - \lambda)/c]/T_n(d/c),$$

where T_n is the Chebyshev polynomial of degree n. The resulting iteration is called Richardson's method. Parameters d and c are the center and focal length respectively ($d \pm c$ are the foci) of a family of ellipses such that one member contains the eigenvalues of A and lies in the right half plane.
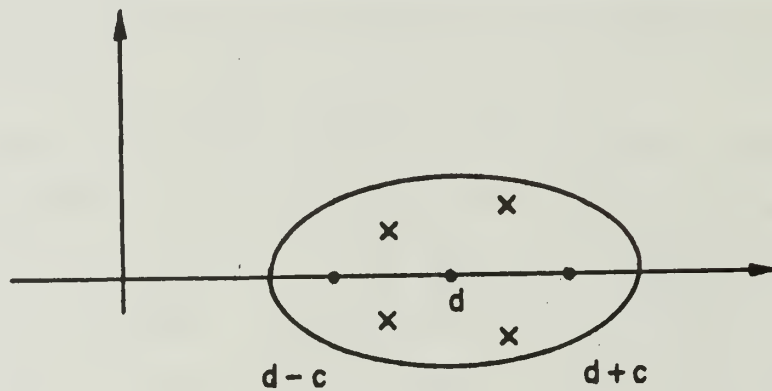

Figure 1. The eigenvalues of A must lie in
an ellipse with center d and foci $d \pm c$
contained in the right (or left) half plane.

The minimax polynomial is that polynomial, subject to (5), whose maximum over any ellipse of the family is less than or equal to the maximum of any such polynomial over that ellipse. The Chebyshev error propagator either is the minimax polynomial or else approximates it with negligible error [9, p. 315].

The Chebyshev propagator depends on the center and focal length of the enclosing ellipses. Optimal values of the center and focal length are determined by the convex hull of the eigenvalues, which may be computed dynamically by the power method, as explained next. (Richardson's method in the form above, (3), is awkward and unstable but convenient for discussion. Another version of the iteration [9, p. 316] is recommended in practice.)

3. Dynamic Estimation of the Eigenvalues. The residual error at step n,

$$r^{(n)} = b - Ax^{(n)} = Ae^{(n)},$$

satisfies

$$r^{(n)} = P_n(A)r^{(0)},$$

analogous to (4) for the true error. For large n, $P_n(\lambda) \doteq [M(\lambda)]^n$ where $M(\lambda) = \{d - \lambda + [(d - \lambda)^2 - c^2]^{1/2}\}/[d + (d^2 - c^2)^{1/2}]$, an approximation which follows from $T_n(w) \doteq e^{n \cosh^{-1} w}/2$. Therefore,

$$r^{(n)} \doteq [M(A)]^n r^{(0)}.$$

Vectors $r^{(n)}$ thus are the iterates of the power method applied to matrix $M(A)$. Suppose μ_1 and $\bar{\mu}_1$ are the dominant eigenvalues of $M(A)$ corresponding to eigenvectors v_1 and $\bar{v}_1$, which are also eigenvectors of A. Then $r^{(n)}$ is approximately a linear combination of v_1 and $\bar{v}_1$ from which μ_1 may be computed. The details of the power method are given in section (5). An eigenvalue, λ_1, of A is the solution of $M_1 = M(\lambda_1)$. An algorithm to compute eigenvalues dynamically and then to compute optimum d and c parameters from the eigenvalues has been developed by Manteuffel [8,9,10] (cf. [7,14]).

4. **The Jacobi Semi-Iterative Method.** The Manteuffel algorithm

estimates optimum d and c parameters from the convex hull of eigenvalues,

which it computes dynamically by the power method. The algorithm may be

applied to any iteration for which the error propagator polynomial trans-

forms into the error propagator polynomial for Richardson's method, if,

also, the iteration yields vectors to which the power method may be ap-

plied. An example is the Jacobi semi-iterative method, described below,

also called the Chebyshev polynomial method [6, p.22;15, p. 355].

Let $A = D - L - U$, where D is a diagonal matrix, L is strictly

lower triangular, and U is strictly upper. The Jacobi method is

$$Dx^{(k+1)} = (L + U)x^{(k)} + b.$$

If $J = D^{-1}(L + U)$,

$$x^{(k+1)} = Jx^{(k)} + s,$$

for $s = D^{-1}b$.

The Jacobi method converges slowly for essentially all practical

problems. To accelerate it, the kth approximation may be taken to be a

linear combination of the first k iterates,

$$y^{(k)} = \sum_{j=0}^{k} a_{kj}x^{(j)}.$$

The coefficients must satisfy the consistency condition

(6)
$$\sum_{j=0}^{k} a_{kj} = 1$$

from which it follows that

$$x - y^{(k)} = \sum_{j=0}^{k} a_{kj}(x - x^{(k)}) = P_k(J)(x - x^{(0)}).$$

The power method cannot be applied using this expression because x is not

known, but the same relation holds for a computable quantity, the residual,

defined by

$$r^{(k+1)} = Jy^{(k)} + s - y^{(k)}.$$

Assume the eigenvalues of J are enclosed in one member of the family of ellipses with center d'and foci d'$\pm$ c' Among all polynomials with coefficients satisfying (6), the minimax polynomial over any such ellipse either is

$$P_k^{(J)}(z') \equiv T_k[(z' - d')/c']/T_k[(1 - d')/c'],$$

or is closely approximated by $P_k^{(J)}$. If $P_k^{(J)}$ is the error propagator polynomial, the iteration is called the Jacobi semi-iterative method (JSI) or the Chebyshev polynomial method. The consistency conditions for the JSI method and Richardson's method respectively are

$$P_k^{(J)}(1) = 1,$$

$$P_k(0) = 1.$$

The error propagators transform into one another as follows. To assume that the eigenvalues of A lie in the right half plane is equivalent to assuming that the eigenvalues of J lie to the left of the vertical line through z' = 1. (Figure 2a). Let E' be an ellipse in the plane left of z' = 1.
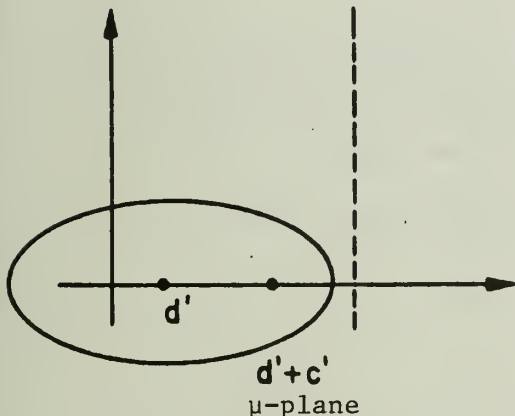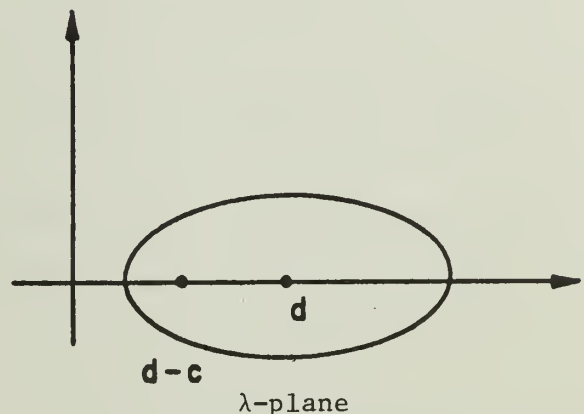

μ-plane

Figure 2a. Jacobi semi-iterative

λ-plane

Figure 2b. Richardson's method

Let $\lambda = 1 - \lambda'$, $d = 1 - d'$ and $c = c'$. Then $d - c > 0$ if $d' + c' < 1$, and

$$P_k^{(J)}(\lambda') = T_k[(\lambda' - d')/c']/T_k[(1 - d')/c']$$

$$= T_k[(d - \lambda)/c]/T_k(d/c)]$$

$$= P_k(\lambda).$$

The best d and c parameters therefore yield the best d' and c' parameters.

The JSI method is a special case of a generalization of Richardson's method attained by the technique of matrix splitting. This generalization is well known but the fact that it includes the JSI method does not seem recognized.

A splitting [2, p. 3] of a matrix A is a decomposition of A into the difference of two matrices,

$$A = M - N,$$

with M such that any system My = w is easy to solve. Richardson's method generalizes to

$$Mx^{(k+1)} = Mx^{(k)} - t_k(Ax^{(k)} - b), \quad k \geq 0.$$

The error at step k obeys the same relation as in equation (4),

$$e^{(k)} = R_{k+1}(M^{-1}A)e^{(0)},$$

where

$$R_{k+1}(\lambda) = \prod_{j=0}^{k} (1 - t_j\lambda).$$

In order to apply the Manteuffel algorithm, the eigenvalues of $M^{-1}A$ must lie in a half plane. An example of a splitting is one obtained by the approximate factorization of the strongly implicit procedure (SIP) [12]. For the JSI method, $M = I$, $N = J$, and $M^{-1}A = I - J$.

5. <u>The Power Method for a Nonsymmetric Matrix.</u> Let B be an

N x N nonsymmetric real matrix with a dominant complex eigenvalue $\dot{\lambda}_1$

corresponding to eigenvector v_1. Complex eigenvalues and eigenvectors

occur in conjugate pairs. Assume for convenience that B is nondefective,

that is, any vector y may be expressed as a sum of eigenvectors,

$$y = \beta_1 v_1 + \bar{\beta}_1 \bar{v}_1 + \dots$$

If $\beta_1 \neq 0$, multiplication by B^k yields a vector,

(7) $$B^k y = \beta_1 \lambda_1^k v_1 + \bar{\beta}_1 \bar{\lambda}_1^k \bar{v}_1 + \dots$$

in which the dominant term is a sum of two eigenvectors. An estimate

of λ_1 may be computed from $B^k y$ and two additional terms, $B^{k+1} y$ and $B^{k+2} y$,

by a method [14, p. 174] which follows from the observation that if c_0 and

c_1 are the coefficients of the polynomial $p_2(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_1) = c_0 + c_1 \lambda + \lambda^2$

then

(8) $$c_0 B^k y + c_1 B^{k+1} y \doteq - B^{k+2} y.$$

Values of c_0 and c_1 that make the left side nearly equal the right are

therefore approximate values of the coefficients of $p_2(\lambda)$. This is an

overdetermined system in two unknowns, c_0 and c_1, a solution of which may

be computed from the method of least squares. For convenience, only the

computation of λ_1 and $\bar{\lambda}_1$ are discussed here. In practice the power method

is used to compute at least four nonreal eigenvalues at a time [8, p. 17].

Let $D = (B^k y, B^{k+1} y)$ be the N x 2 matrix of system (8). Thus,

$$D \binom{c_0}{c_1} = - B^{k+2} y.$$

The least squares solution is the solution of the 2 x 2 system

$$D^T D \binom{c_0}{c_1} = - D^T B^{k+2} y.$$

Four inner products are required to compute $D^T D$. The number of multiplicatons is $4(N-1)$.

      The discussion of the power method in this section has been simplified by considering only two nonreal eigenvalues. In practice the power method is used to compute at least four nonreal eigenvalues at a time [10, p. 18].

      6.   <u>A Reduction in the Work to Solve the Least Squares Problem</u>. A subsystem of equations could be selected from the overdetermined system (8) at random, and the method of least squares applied to the subsystem to reduce the work, but the solution must not be grossly perturbed. In this instance the number, N, of equations, or statistical observations, is so much greater than the number, 2, of parameters, c_0 and c_1, that one can reasonably expect variations about the true value of the parameters to be uniform for a smaller number of observations than N. This is an intuitive reason why the method of least squares could be applied to a subsystem. A closer look may be more convincing.

      Let ν_i be the ith component of v_1; let η_i, $\eta_i{'}$, and $\eta_i{''}$ be the ith components of $B^k y$, $B^{k+1} y$, and $B^{k+2} y$ respectively. From the equations

$$\beta_1 \lambda_1{}^k v_1 + \bar{\beta}_1 \bar{\lambda}_1{}^k \bar{v}_1 + \ldots = B^k y$$

$$\beta_1 \lambda_1{}^{k+1} v_1 + \bar{\beta}_1 \bar{\lambda}_1{}^{k+1} \bar{v}_1 + \ldots = B^{k+1} y$$

$$\beta_1 \lambda_1{}^{k+2} v_1 + \bar{\beta}_1 \bar{\lambda}_1{}^{k+2} \bar{v}_1 + \ldots = B^{k+2} y$$

it follows that

$$\lambda_1{}^k \beta_1 \nu_i + \bar{\lambda}_1{}^k \bar{\beta}_1 \bar{\nu}_i + \zeta_i = \eta_i$$

$$\lambda_1{}^{k+1} \beta_1 \nu_i + \bar{\lambda}_1{}^{k+1} \bar{\beta}_1 \bar{\nu}_i + \zeta_i{'} = \eta_i{'}$$

(9) $$\lambda_1{}^{k+2} \beta_1 \nu_i + \bar{\lambda}_1{}^{k+2} \bar{\beta}_1 \bar{\nu}_i + \zeta_i{''} = \eta_i{''}$$

where ζ_i, ζ_i', and ζ_i'' are the contributions due to low order

terms. If c_0 and c_1 are, as before, the coefficients of

$$p_2(\lambda) = (\lambda - \lambda_1)(\lambda - \bar{\lambda}_1) = c_0 + c_1\lambda + \lambda^2,$$

then

$$c_0 n_i + c_1 n_i' + n_i'' = \zeta_i + \zeta_i' + \zeta_i''.$$

Any two equations uniquely determine c_0 and c_1,

$$c_0 n_i + c_1 n_i' + n_i'' = \zeta_i + \zeta_i' + \zeta_i''$$

$$c_0 n_j + c_1 n_j' + n_j'' = \zeta_j + \zeta_j' + \zeta_j''$$

If the right sides, due to low order terms, are small, then the solution

of

(10)

$$\tilde{c}_0 n_i + \tilde{c}_1 n_i' + n_i'' = 0$$

$$\tilde{c}_0 n_j + \tilde{c}_1 n_j' + n_j'' = 0$$

approximates (c_0, c_1). Even though the low order terms are small in

$B^k y$, they are not necessarily small component by component. It is a pre-

caution against this to include more equations in the system (10) deter-

mining $\tilde{c}_0$ and $\tilde{c}_1$. Since $B^k y$ is nearly $\beta_1 \lambda_1^k v_1 + \bar{\beta}_1 \bar{\lambda}_1^k \bar{v}_1$ components of the

low order terms are generally small, although, of course, it may happen

that only one component of $\beta_1 \lambda_1^k v_1 + \bar{\beta}_1 \bar{\lambda}_1^k \bar{v}_1$ is dominant in $B^k y$. In other

words, a system

$$\tilde{c}_0 n_{i_1} + \tilde{c}_1 n_{i_1}' = -n_{i_1}''$$
$$\vdots \qquad \vdots$$
$$\tilde{c}_0 n_{i_\ell} + \tilde{c}_1 n_{i_\ell}' = -n_{i_\ell}''$$

of fewer than N equations would probably yield a satisfactory least squares

solution. With a smaller system, the storage for the power method is also reduced.

7. **A Test for the Size of the Subsystem.** Whether the number of
equations is adequate may be determined in a precise way. If there are no
low order terms the column vectors $u = (\eta_{i_1}, \ldots, \eta_{i_\ell})^T$, $u' = (\eta'_{i_1}, \ldots, \eta'_{i_\ell})^T$,
and $u'' = (\eta''_{i_1}, \ldots, \eta''_{i_\ell})^T$ are linear dependent. Therefore the matrix

$$(11) \qquad C = (u \ u' \ u'')$$

has a singular value zero. In reality, the columns are linearly in-
dependent with a nonzero singular value of smallest magnitude. The larger
this singular value is, the more the lower order terms dominate in (9).
If the system is small, it would be economical to compute the singular
values as a check on accuracy. The work of computing the singular values
could be applied to the solution of the least squares problem, a savings
which is explained in the next section.

8. **The Singular Value Decomposition.** Any $m \times n$ matrix, C,
may be decomposed in the form

$$C = U \ \Sigma \ V$$

where U is an $m \times m$ orthogonal matrix, Σ is an $m \times n$ diagonal matrix,
the entries of which are the singular values, and V is an $n \times n$ orthogonal
matrix [4, 11]. This is called the singular value decomposition (SVD).
The ratio of the largest (in magnitude) singular value to the smallest
is a measure of the linear independence of the column vectors of C.
The SVD reduces the question of linear independence to a single quantity,
the ratio of the extreme singular values.

Let $C = (u \ u' \ u'')$ be the $\ell \times 3$ matrix (11) resulting from
the power method. The steps to compute the SVD form two units. First,
C is reduced to bidiagonal form in which the only nonzero elements lie on

the main diagonal and the superdiagonal (just above the main diagonal) [4, 3, 11]. A modification of the QR algorithm, due to Golub [11, p. 387; 5], may then be used to reduce the bidiagonal matrix to diagonal form. (Other methods are possible but this one is stable and widely accepted.)

Reduction of C to bidiagonal form may be made by Householder transformations. A Householder transformation is a symmetric and orthogonal matrix of the form $P = I - 2ww^T$, where $w^Tw = 1$, which may be constructed in such a way as to transform a given vector z, whose first component must be nonzero, into Pz, each component of which, except the first, is zero. The formulas for P are:

$$\sigma = z^Tz;$$

$$s = z + \sigma e_1, \text{ where } e_1^T = (1,0,\ldots,0);$$

$$\beta = s^Ts/2;$$

$$P = I - ss^T/\beta = I - \frac{s}{\sqrt{\beta}}\frac{s^T}{\sqrt{\beta}}.$$

They yield

$$Pz = \sigma e_1.$$

More details may be found in the text of Stewart [11, p. 230]. If <u>operation</u> means either a multiplication or a division, 4ℓ operations are required for σ, β, and division of s by $\beta^{1/2}$ in the formula for P.

The actual reduction to bidiagonal form is by a sequence of Householder transformations multiplying C on the left and on the right. Those on the left eliminate nonzero elements under the diagonal, whereas those on the right eliminate nonzero elements above the superdiagonal. Let P_0, P_1, P_2 denote the transformations on the left and Q_0 the transformation on the right. These yield, where x denotes a possibly nonzero element:

$$P_0 C = \begin{pmatrix} x & x & x \\ 0 & x & x \\ \vdots & \vdots & \vdots \\ 0 & x & x \end{pmatrix}$$

$$C_2 = P_1 P_0 C = \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ \vdots & \vdots & \vdots \\ 0 & 0 & x \end{pmatrix}$$

$$C_3 = P_2 P_1 P_0 C = \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$C_4 = P_2 P_1 P_0 C Q_0 = \begin{pmatrix} x & x & 0 \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{pmatrix} .$$

In order to use the results of the reduction for the solution of the least squares problem, the matrix, D, of which consists of the first two columns of C, it is convenient not to multiply by Q_0 before multiplication by P_0 and P_1. The first two columns of $C_2 = P_1 P_0 C$ will then be the least squares matrix, D, reduced to upper triangular form, from which the solution of the least squares problem may be computed at no expense.

Construction of Q_0 and multiplication on the right by Q_0 are negligible since Q_0 is a 3 x 3 matrix. Construction and multiplication on the left by P_0, P_1, and P_2 take 10ℓ, 8ℓ, and 6ℓ operations respectively. The total number of operations to compute C_4 is 24ℓ.

The further reduction of C_4 to diagonal form is not discussed.

9. <u>Conclusion</u>. Richardson's method and a generalization of Richardson's method that includes the Jacobi semi-iterative method employ Chebyshev acceleration parameters determined by the convex hull of eigenvalues of the error propagator matrix. The Manteuffel algorithm, which estimates parameters dynamically, performs two basic operations: It computes the convex hull by the power method, which requires a least squares procedure; and it computes the best parameters from the convex hull. Computation of the best parameters is inexpensive once the eigenvalues are known. The cost of the Manteuffel algorithm is in the solution of the least squares problem. Examination of a method to reduce execution time and storage requirements for the least squares problem by using fewer components of $B^k y$ shows that the method works if the columns of matrix C, (11), are linearly dependent. The singular values of C may be used to test linear dependence and at the same time provide the solution of the least squares problem. This summarizes the report through the last section.

If fewer components of $B^k y$ are used, there are fewer rows of C and the singular values are inexpensive to compute. Singular values have allowed the use of fewer components, but now fewer components allow the free use of singular values. Inexpensive singular values could be computed after each iteration to test whether new eigenvalues and therefore improved acceleration parameters are ready. Testing singular values displaces the old strategy of waiting a fixed number of iterations before improving acceleration parameters. The Manteuffel algorithm now falls in step with the iterative method further enhancing convergence.

REFERENCES

1. P. A. Businger and G. H. Golub, Algorithm 358, Singular value de-
    composition of a complex matrix, Comm. ACM 12, 1969, pp. 564-565.

2. Paul Concus, Gene H. Golub, and Dianne P. O'Leary, A generalized
    conjugate gradient method for the numerical solution of elliptic
    partial differential equations, in Sparse Matrix Computations,
    edited by James R. Bunch and Donald J. Rose, Academic Press,
    New York, 1976.

3. George Forsythe, Michael A. Malcolm, Cleve B. Moler, Computer
    Methods for Mathematical Computations, Prentice-Hall, Englewood
    Cliffs, New Jersey 07632, 1977.

4. G. Golub and W. Kahan, Calculating the singular values and pseudo-
    inverse of a matrix, J. SIAM Numer. Anal., Ser. B, Vol. 2, No. 2,
    1965.

5. G. H. Golub and C. Reinsch, Singular value decompositions and least
    squares solutions, Numer. Math. 14, 1970, pp. 403-420.

6. L. W. Hageman, The Estimation of Acceleration Parameters for the
    Chebyshev Polynomial and the Successive Overrelaxation Iteration
    Methods, Memorandum prepared for USAEC by Westinghouse Electric
    Corporation, June, 1972, Available from National Technical Information
    Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield,
    VA 22151.

7. G. Kjellberg, On the convergence of successive over relaxation applied
    to a class of linear systems of equations with complex eigenvalues,
    Ericsson Technics, No. 2, 1958, pp. 245-258.

8. T. A. Manteuffel, An Iterative Method for Solving Nonsymmetric
    Linear Systems with Dynamic Estimation of Parameters, Rep. UIUCDCS-R-75-
    758, University of Illinois, U-C, Urbana, Illinois 61801, Oct., 1975.

9. T. A. Manteuffel, The Tchebyshev iteration for nonsymmetric linear
    systems, Numer. Math. 28, 1977, pp. 307-327.

10. T. A. Manteuffel, Adaptive Procedure for Estimating Parameters for the
    Nonsymmetric Tchebyshev Iteration, Sandia Laboratories Report SAND77-8239,
    Livermore, CA, June 1977.

11. G. W. Stewart, Introduction to Matrix Computation, Academic Press,
    New York, 1973.

12. H. L. Stone, Iterative solution of implicit approximations of multi-
    dimensional partial differential equations, SIAM J. Numer. Anal.,
    5 (1968), pp. 530-558.

13. J. H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965.

14. H. E. Wrigley, Accelerating the Jacobi method for solving simultaneous equations by Chebyshev extrapolation when the eigenvalues of the iteration matrix are complex, Computer Journal, Vol. 6, July, 1963, pp. 169-176.

15. David M. Young, Iterative Solution of Large Linear Systems, Academic Press, New York, New York, 1971.

Title and Subtitle		5. Report Date April 1978
USE OF THE SINGULAR VALUE DECOMPOSITION TO INCREASE EXECUTION AND STORAGE EFFICIENCY OF THE MANTEUFFEL ALGORITHM FOR THE SOLUTION OF NONSYMMETRIC LINEAR SYSTEMS		6.

Author(s) Paul E. Saylor	8. Performing Organization Rept. No.

Performing Organization Name and Address Department of Computer Science University of Illinois at Urbana-Champaign Urbana, IL 61801	10. Project/Task/Work Unit No.
	11. Contract/Grant No. NSF 76-81100

Supplementary Notes

Abstracts Optimum Chebyshev parameters may be computed dynamically by the Manteuffel algorithm for use with a generalization of Richardson's iterative method and the Jacobi semi-iterative method to solve nonsymmetric linear systems. Parameters are determined by the algorithm from the convex hull of the eigenvalues of the error propagator, which the algorithm also computes by the power method. The singular value decomposition may be used to test the reliability of a simple technique to reduce execution and storage costs of the power method. The same technique also yields an inexpensive singular value decomposition making it feasible to determine precisely when to call the Manteuffel algorithm.

Key Words and Document Analysis. 17a. Descriptors

linear algebraic equations, iteration, eigenvalues, eigenvectors, numerical analysis

Identifiers/Open-Ended Terms

COSATI Field/Group

Availability Statement	19. Security Class (This Report) UNCLASSIFIED	21. No. of Pages
	20. Security Class (This Page) UNCLASSIFIED	22. Price